Assessments, assessments, assessments: notes from the 2023 CMS summer meeting.

Kseniya Garaschuk (University of the Fraser Valley)

Education Notes bring mathematical and educational ideas forth to the CMS readership in a manner that promotes discussion of relevant topics including research, activities, issues, and noteworthy news items. Comments, suggestions, and submissions are welcome.

Kseniya Garaschuk, *University of the Fraser Valley (kseniya.garaschuk@ufv.ca)*
John Grant McLoughlin, *University of New Brunswick (johngm@unb.ca)*

As I was going through my office reorganizing it for my sabbatical, I stumbled upon my notes from the education sessions of the 2023 CMS Summer meeting. So many great ideas, all in one place. Good thing I'm forced into occasional organizing.

One particular theme stood out to me at the 2023 CMS Summer meeting (and selfishly the one theme that I never get tired of discussing). So in this piece, I highlight a few presentations with a general common theme of assessments, specifically the different ways we can think about, approach and implement assessments. This article also has a second, more subtle goal: to give readers a taste of education sessions at CMS meetings. As such, below I summarize my thoughts from a few talks, making for a mixture of presented materials and my take on it.

*****

Assessment aftermath or notes around Darja Barr's talk

As a math instructor, you have surely thought of test anxiety before. Maybe you experienced it yourself, as a student or as an educator — I myself nearly had a nervous breakdown before my PhD defense and to this day I get nervous when my students write a test. Of course tests are stressful by nature. Research shows that the higher the person's anxiety, the lower their performance on the test. However, in contrast to the actual research, we often hear the argument that higher anxiety is caused by lower preparation, so the anxiety is allegedly the middle man in this lower-preparation-lower-result game.

Not all tests are created equal. Maybe we can affect the stress levels by varying the types of assessment we give our students, for example in terms of the grade weight — high stakes versus low stakes. Sadly, each format has their pros and cons. Specifically, while low stakes tests are, in fact, low stakes and hence less stressful, they also produce a 'delayed consequences' phenomenon. This occurs when students don't consider it a big deal to fail a small test since it isn't worth much and hope to do better on the next one. Unfortunately, despite the fact that students optimistically believe that future-them is always better at math than present-them, hoping to do better is not the same as planning to do better.

While I can see many flaws with traditional testing, the assessments' transitory fleeting nature is the most significant. As an instructor, I see any mid-semester assessment as a learning opportunity; in cooking terms, it's a chance to taste the dish during the cooking process and adjust spices/temperature/consistency as necessary. However, without explicit prompting, many students do not use the test results to assess the situation, to reflect on their mathematical mistakes and adjust their approach to studying for the course. This very issue was one of the big motivators in my own research on collaborative exams. I have also used a variety of test-aftermath assignments, where students are asked to not only identify and correct their mistakes, but also go through the test and for each problem find at least 3 examples from the available course materials (lecture notes, homeworks, practice problems, quizzes, past tests) that target the same concept. So we have established that they have access to the necessary materials, but how do I know if they are using them? To allow for honest self-assessment and to prompt students to look inwards at their study habits, Darja uses a "What did I do?" survey to give students a tangible way to estimate their engagement in the course. Here is one version of it:

Post-Test Behaviour Self-Evaluation

Give yourself points for each action item:

1. *Attendance*

   —*Attended every lecture*   *(2 points)*

   —*Missed a single lecture*   *(1 points)*

— *Missed more than one lecture* *(0 points)*

2. *Practice*

  — *Complete the assigned problems from each section in the text* *(2 points)*

  — *Attempt most of the problems* *(1 points)*

  — *Attempt some or none of the problems* *(0 points)*

3. *Assignments*

  — *Completed every assignment (grade not important)* *(2 points)*

  — *Missed completing one assignment* *(1 points)*

  — *Missed completing more than one assignment* *(0 points)*

4. *Tutorials*

  — *Attended every tutorial and completed the work* *(2 points)*

  — *Missed a single tutorial* *(1 points)*

  — *Missed more than one tutorial* *(0 points)*

5. *Seeking Help*

  — *Emailed your prof with a MATH question at least twice* *(2 points)*

  — *Emailed your prof with a MATH question once* *(1 points)*

  — *Did not email your prof with any math questions* *(0 points)*

As the prof, you can decide on your minimum cut-off, include more categories for the point-worthy behaviour or exclude some that are there. But this type of survey gives a quick and easy tool for students to assess their approach to the class and think about how they position themselves for success. Furthermore, embedded within this survey are clear suggestions as to how a student could course-correct since categories provide proactive ways to engage with the class.

*****

Communicating math or notes around Fabian Parsch's talk

Any mathematics course is a communications course. A quote attributed to Einstein eloquently summarizes the idea: "You don't really understand something unless you can explain it to your grandmother". The true study of the subject is about understanding and applying various mathematical concepts, which is achieved and strengthened by students' learning to communicate these concepts to themselves and others. In my practice, I have tried a variety of things, including for example learning journals.

Fabian's context was an engineering calculus course where the instructional team focused on developing students' communication skills with motivation that, as future engineers, the students will need to communicate technical aspects to their non-technical clients. Best practice for any kind of assignments includes a draft stage, which can provide an opportunity for the student to get feedback on their work so far; once again cooking-wise, this is akin to someone else tasting your dish during preparation and making suggestions for possible adjustments. Ability to get feedback and act on it is especially important for writing assignments in technical courses — a skill that students likely have limited (if any) experience with, so building in the draft feedback cycle is a helpful step in the process. Next up is marking: grading writing assignments can be an intimidating and daunting task, so a rubric to guide both students and markers is a must and so is the process of sharing and discussing such rubric with students well before the drafts are due. This is Fabian's rubric:

	2 points	1 point	0 points
Audience	The answer could be understood by a student of this course without the need to ask questions of clarification, even if they never saw this assignment before.	The answer could mostly be understood by a student of this course, but issues persist. For example, questions of clarification would be necessary.	The answer could only be understood by someone with extensive knowledge of the subject (e.g. a TA or an instructor).
Style	The answer uses full sentences that guide the reader through the steps.	Full sentences are being used, but to limited extent and/or in isolation.	The answer is mostly a "wall of formulas" with no explanation.
Vocabulary	Accurate usage of terminology, used appropriately in context	Terminology is generally used, but is either lacking or used excessively	Vague, weak or no usage of appropriate terminology

There are many ways of developing a rubric. For instance, instead of directly providing one, you can provide examples of sample submissions and build a rubric together with students based on the discussions of what they found good or bad, helpful or frustrating, useful or obstructive in the writing samples.

There are many ways to design writing assignments as well. With the ever-improving quality of AI-generated text, we need to get a bit more creative ourselves. Here are some ideas:

- a memo on the results of a technical paper or report
- an executive summary of a technical presentation
- an explanation or critique of a piece of math found in the media this week/month

*****

Not all tests need to be written or notes around Diana Skrzydlo's talk

One of the biggest noticeable impacts of the pandemic on education was our collective re-thinking of assessment practices. With invigilated exams no longer an option, we were forced to question everything about the state of the exams, from their goals down to the format. Veselin Jungic and I wrote about our view of assessments and academic integrity in the times of COVID, but I'm sad to say that I see more and more people going "back to normal" in their teaching and not capitalizing on the lessons learned in early days of the lockdown. In particular, I am on the lookout more than ever before for more meaningful formative and summative assessment formats that underline more authentic and purposeful teaching (and, if possible, also lighten the academic integrity concerns). Enter oral exams.

Oral exam is a dialogue. It is an opportunity for students to explain conceptual ideas in words, while allowing instructors to ask for clarification. Just like any other assessment though, it needs to align with the course goals and in-term opportunities to practice oral communication. It is essential to create clear guidelines for what will be asked of the student and how the examiner will assess it, since few (if any) students come in with previous experience in this format and hence find it very intimidating.

Diana's format is as follows: 15 minutes, 5 questions each of a different type and drawing on a different section or topic of a course. The 5 questions types are:

- definition,
- advantages/disadvantages,
- compare/contrast,
- describe a process,
- predict an impact.

Here are some specific examples from her statistics course:

- Compare and contrast: Binomial vs Poisson distribution, Type I vs Type II error.
- Describe a procedure: joint transformation, Bayesian estimation, generating random walk/Brownian Motion.

For practice, students are encouraged to think of at least one of each type of question for each section of the material. Note that with this approach, the questions target different levels of Bloom's taxonomy: from remembering and applying in the definitions and describing procedures to evaluating and creating in compare/contrast and predicting impact categories.

*****

Some of the most inspirational are my notes and thoughts from the session run by Peter Taylor and Chris Suurtamm that involved a panel with high school teachers. While other sessions largely focused on specific teaching tools and techniques, this session focused on overall approach and philosophy of teaching. I often think about similarities and differences in teaching math and English in a standard first year university class. Both are service courses meaning the audience consists of mostly non-majors in the corresponding subject, both are mandatory classes that students often do not wish to be enrolled in, both appear to students as a necessary evil that isn't relevant to their future careers or current interests, both carry the reputation of "I was never good at...", both a mathematical solution and a well-organized essay are just a series of logical steps presented within the specified context, both aim to teach communication and general critical

thinking skills rather than being heavy in the discipline-specific jargon. However, take a look at the course outlines. English courses give outcomes, not context: they mention types of analysis or writing students will do (rhetoric analysis, close reading, response and reflection writing), but not what specific works they will use in the course (Shakespeare vs Tolstoy). Math courses give content, not outcomes: students will be able to apply differentiation rules, row-reduce a matrix, calculate probability distributions and so on with no mention of neither communicating their findings nor developing higher-level problem solving skills, analytical and computational thinking that demonstrate the power and relevance of mathematics to mathematicians and non-mathematicians alike. Of course the two subjects of math and English differ by nature, but contrasting the two helps me focus on the ultimate goal stated by someone in the session: we need to be teaching more than the alphabet of math.

*****

What you see above is a snapshot of just a small portion of the talks. I took many more notes about many other topics. If you've read this far, there are two "morals" I would like to get across. First of all, go to the education sessions at the CMS meetings. If teaching is a part of your everyday life, you will find many inspirational and very practical ideas in these education sessions to bring into your own classroom. Secondly, it was August, two months since the conference and this was the first time I looked through my notes from the session. I am generally a good note taker, so the notes were easy to follow as I re-discovered some ideas that I thought were great back in June. Clearly, I need to become more systematic about reviewing notes I take for future development of my teaching practice. But also in general, as a community, we need to do better with sharing and archiving this information. For people who couldn't make it to the conference or couldn't make it to a particular talk, how can CMS collect and make available the materials and ideas that are being developed by educators across Canada that we can all benefit from? My first attempt was writing this piece; if you have other ideas, please contact me – I'm happy to talk more.